

Public Dataset 2001 Patient Discharge Data

Our goal in the development of a revised patient-level dataset is to maintain the value of the data elements currently available to the public while minimizing the risk to patient confidentiality. The interim public patient-level dataset includes patient zip code, demographic variables and clinical information (see Table 1). The interim dataset would be released without an extensive review process.

OSHDPD has made every effort to retain many of the data elements that were available on public datasets, including most of the demographic variables. We will also include, length of stay and admit quarter. We are continuing to explore the creation of public data products which will include Record Linkage Number (RLN).

We plan to employ a masking routine to protect the confidentiality of these files. Records that have unique combinations of demographic and patient geographic variables would have one or more of those variable values blanked out to ensure the file is de-identified. Masking of the demographic data elements in this dataset would occur in a predetermined order (see Table 1). Age in years at admission would be masked first and patient zip code masked last.

As indicated in Table 1, the interim dataset includes age in years as well as two pre-defined sets of age categories. Furthermore, patient county for counties with less than 30,000 population would be aggregated into three groupings as described on Table 1.

Table 2 is a comparison of the data elements contained in the former public Versions A and B with the current public dataset.

Table 1

Standard File 2001 Patient Discharge Data	
<i>Data Elements Subject to Masking and Masking Order</i>	
Age in years Ethnicity Race Sex Age Category (20)* Age Category (5)* Small County (if applicable)** Admit Quarter Patient 5-digit Zip Code Patient 3-digit Zip Code OSHDP ID	
<i>Additional Data Elements (not masked)</i>	
Condition Present on Admission Days from Admit to Procedure Do Not Resuscitate Order DRG Expected Source of Payment Length of Stay MDC Other E-Codes Patient Disposition Principal Diagnosis Principal External Source of Injury Codes (E-Codes) Principal Procedure Secondary Diagnoses (24) Secondary Procedures (20) Source of Admission Total Charges Type (level) of Care Type of Admission	

* Age categories:

20 = < 1 year; 5-year increments to 84 (e.g., 1-4, 5-9, 10-14...); 85+

5 = < 1 year; 1-17 years; 18-34; 35-64; 65+

**If a county has < 30,000 population, patient county will be grouped with another county or counties and assigned to one of three values: NE (Modoc, Plumas and Sierra), NW (Colusa, Glenn and Trinity), and CE (Alpine, Inyo, Mariposa and Mono).